

DETERMINING A TANGENT SPACE AND FILTERING DATA ONTO A
MANIFOLD

ABSTRACT OF THE DISCLOSURE

5

A technique for determining the number of constraints on a set of input data, or equivalently the topological dimension, especially when such data are produced by a nonlinear system, such as a pathological vocal system or econometric data and the like.

The technique characterizes the tangent space about a predetermined base point by

10 identifying a maximal set of non-redundant nonlinear fits to the data. It needs only a modest number of data points and does not assume prior knowledge of the functional form of the true constraints, other than smoothness. Each fit is equivalent to a set of contours (including curves, surfaces, and other manifolds), with the data themselves all lying along the zero-value contour of the fit. For each fit, the gradient of the fit at the

15 base point in the uphill direction across the contours identifies the constraint direction. Considering all fits simultaneously, the number of constraint directions that are linearly independent provides the number of constraints in the neighborhood of the base point. The remaining unconstrained directions define the tangent space, and its dimensionality, which is precisely the number of linearly independent unconstrained directions, is

20 precisely the inferred topological dimensionality of the original data.